Web-Only Mathematical Appendix for "Trade Shocks and Labor Adjustment: A Structural Empirical Approach"

by Erhan Artuç, Shubham Chaudhuri and John McLaren

I. Appendix 1: Derivation of Equilibrium Conditions with the Extreme Value Distribution.

A. *Overview of the Derivation.*

The cumulative distribution function for the extreme value distribution with zero mean is given by:

$$F(\varepsilon) = \exp(-\exp(-\varepsilon/\nu - \gamma)),$$

where $\gamma \cong 0.5772$ is Euler's constant. The associated density function is:

$$f(\varepsilon) = (1/\nu)\exp(-\varepsilon/\nu - \gamma - \exp(-\varepsilon/\nu - \gamma)).$$

In the following subsection we will derive equation (2), which relates gross flow rates to the value function. In the subsection after that we will derive the form for the option-value function reported in (3):

(1) $$\bar{\varepsilon}_t^{ij} \equiv \beta E_t[V^j(L_{t+1}, s_{t+1}) - V^i(L_{t+1}, s_{t+1})] - C^{ij},$$

(2) $$\bar{\varepsilon}_t^{ij} \equiv \beta E_t[V_{t+1}^j - V_{t+1}^i] - C^{ij} = \nu[\ln m_t^{ij} - \ln m_t^{ii}],$$

and

(3) $$\Omega(\bar{\varepsilon}_t^i) = -\nu \ln m_t^{ii}.$$

B. *The m^{ij} function.*

The gross flow of workers from i to j at date t, m_t^{ij}, is equal to the probability that a given i-worker will switch to j at date t, or the probability that, for an i-worker, utility $w_t^i + \varepsilon_t^j + \beta E_t[V^j(L_{t+1}, s_{t+1})] - C^{ij}$ will be higher for a move to j than for any of the other $n-1$ options. In other words, from (1),

$$m_t^{ij} = Prob_{\varepsilon_t}\left[\bar{\varepsilon}_t^{ij} + \varepsilon_t^j \geq \bar{\varepsilon}_t^{ik} + \varepsilon_t^k \text{ for } k = 1, \ldots, n\right].$$

Suppressing the time subscript, this can be written:

$$m^{ij} = \int_{-\infty}^{\infty} f(\varepsilon^j) \prod_{k \neq j} F(\varepsilon^j + \bar{\varepsilon}^{ij} - \bar{\varepsilon}^{ik}) d\varepsilon^j.$$

Define, for convenience: $x \equiv \varepsilon^j/\nu + \gamma$, $z^j \equiv \bar{\varepsilon}^{ij}$, $\bar{\varepsilon}^{ik} = z^k$, and $\lambda \equiv \log(\frac{\sum_{k=1}^n \exp(z^k/\nu)}{\exp(z^k/\nu)})$. Then the expression for gross flows can be rewritten:

$$\begin{aligned}
m^{ij} &= \frac{1}{\nu}\int \exp(-\varepsilon^j/\nu - \gamma\ - \exp(-\varepsilon^j/\nu - \gamma))\prod_{k\neq j}\exp(-\exp(-[\varepsilon^j + \overline{\varepsilon}^{ij} - \overline{\varepsilon}^{ik}]/\nu - \gamma))d\varepsilon^j \\
&= \frac{1}{\nu}\int \exp(-\varepsilon^j/\nu - \gamma\ - \exp(-\varepsilon^j/\nu - \gamma))\exp(-\sum_{k\neq j}\exp(-[\varepsilon^j + \overline{\varepsilon}^{ij} - \overline{\varepsilon}^{ik}]/\nu - \gamma))d\varepsilon^j \\
&= \frac{1}{\nu}\int \exp(-\varepsilon^j/\nu - \gamma)\exp(-\sum_{k=1}^{n}\exp(-[\varepsilon^j + \overline{\varepsilon}^{ij} - \overline{\varepsilon}^{ik}]/\nu - \gamma))d\varepsilon^j \\
&= \frac{1}{\nu}\int \exp\left[(-\varepsilon^j/\nu - \gamma) - \sum_{k=1}^{n}\exp(-[\varepsilon^j + \overline{\varepsilon}^{ij} - \overline{\varepsilon}^{ik}]/\nu - \gamma)\right]d\varepsilon^j \\
&= \frac{1}{\nu}\int \exp\left[(-\varepsilon^j/\nu - \gamma) - \exp((-\varepsilon^j/\nu - \gamma))\sum_{k=1}^{n}\exp(-[z^j - z^k]/\nu)\right]d\varepsilon^j \\
&= \frac{1}{\nu}\int \exp\left[(-\varepsilon^j/\nu - \gamma) - \exp((-\varepsilon^j/\nu - \gamma))\left(\sum_{k=1}^{n}\exp(z^k/\nu)\right)/\exp(z^j/\nu)\right]d\varepsilon^j \\
&= \int \exp(-x - \exp(-(x-\lambda)))dx.
\end{aligned}$$

This again can be rewritten:

$$m^{ij} = \exp(-\lambda)\int \exp(-(x-\lambda) - \exp(-(x-\lambda)))dx.$$

Now set $y = x - \lambda$. Noting that the antiderivative of

$$\exp(-y - \exp(-y))$$

is

$$\exp(-\exp(-y)),$$

we can derive:

$$\begin{aligned}
m^{ij} &= \exp(-\lambda)\int \exp(-y - \exp(-y))dy \\
&= \exp(-\lambda) \\
&= \frac{\exp(z^j/\nu)}{\sum_{k=1}^{n}\exp(z^k/\nu)} \\
&= \frac{\exp(\overline{\varepsilon}^{ij}/\nu)}{\sum_{k=1}^{n}\exp(\overline{\varepsilon}^{ik}/\nu)}.
\end{aligned}$$

Given that $\overline{\varepsilon}^{ii} \equiv 0$, this yields (2).

C. The Option-Value Function.

Define:

$$
\begin{aligned}
\Psi^{ij} &\equiv \int_{-\infty}^{\infty}(\varepsilon^j - C^{ij}) f(\varepsilon^j) \prod_{j\neq k} F(\varepsilon^j + \overline{\varepsilon}^{ij} - \overline{\varepsilon}^{ik}) d\varepsilon^j \\
&= \frac{1}{\nu}\int(\varepsilon^j - C^{ij}) \exp(-\varepsilon^j/\nu - \gamma - \exp(-\varepsilon^j/\nu - \gamma)) \prod_{k\neq j} \exp(-\exp(-[\varepsilon^j + \overline{\varepsilon}^{ij} - \overline{\varepsilon}^{ik}]/\nu - \gamma))d\varepsilon^j
\end{aligned}
$$

Going through the steps of Subsection (B), we find:

$$
\begin{aligned}
\Psi^{ij} &= \int(\nu(x-\gamma) - C^{ij}) \exp(-x - \exp(-(x-\lambda)))dx \\
&= (-C^{ij} - \nu\gamma) \exp(-\lambda) + \nu \int x \exp(-x - \exp(-(x-\lambda)))dx \\
&= (-C^{ij} - \nu\gamma) \exp(-\lambda) + \nu \exp(-\lambda) \int x \exp(-x + \lambda - \exp(-(x-\lambda)))dx
\end{aligned}
$$

We know that $\exp(-\lambda) = m^{ij}$ from the previous derivation. Substituting this in:

$$
\begin{aligned}
\Psi^{ij} &= (-C^{ij} - \nu\gamma)m^{ij} + \nu m^{ij} \int x \exp(-x + \lambda - \exp(-(x-\lambda)))dx \\
&= (-C^{ij} - \nu\gamma)m^{ij} + \nu m^{ij} \int x \exp(-x + \lambda - \exp(-(x-\lambda)))dx \\
&\quad + \nu m^{ij} \int \lambda \exp(-x + \lambda - \exp(-(x-\lambda)))dx \\
&\quad - \nu m^{ij} \int \lambda \exp(-x + \lambda - \exp(-(x-\lambda)))dx \\
&= (-C^{ij} - \nu\gamma)m^{ij} + \nu m^{ij} \int (x-\lambda) \exp(-x + \lambda - \exp(-(x-\lambda)))dx \\
&\quad + \nu m^{ij} \int \lambda \exp(-x + \lambda - \exp(-(x-\lambda)))dx \\
&= (-C^{ij} - \nu\gamma)m^{ij} + \nu m^{ij} \int y \exp(-y - \exp(-y))dy + \nu\lambda m^{ij} \int \exp(-y - \exp(-y))dy \\
&= (-C^{ij} - \nu\gamma)m^{ij} + \nu m^{ij} \int y \exp(-y - \exp(-y))dy + \nu\lambda m^{ij}.
\end{aligned}
$$

Noting that $\int y \exp(-y - \exp(-y))dy = \gamma$ (Euler's constant) (Patel, Kapadia and Owen (1976, p. 35)), we can simplify:

$$
\begin{aligned}
\Psi^{ij} &= (-C^{ij} - \nu\gamma)m^{ij} + \nu\lambda m^{ij} + \nu\gamma m^{ij} \\
&= -C^{ij} m^{ij} - \nu \log(m^{ij}) m^{ij} \\
&= m^{ij}\left(-C^{ij} - \nu \log(m^{ij})\right).
\end{aligned}
$$

Adding this up across possible destinations j, note that the utility of a worker in i is equal to:

$$\begin{aligned} V_t^i &= w_t^i + \sum_{j=1}^{n} \left(\Psi_t^{ij} + \beta m_t^{ij} V_{t+1}^j \right) \\ &= w_t^i + \sum_{j=1}^{n} \left[m_t^{ij} (-\nu \log(m_t^{ij}) - C^{ij} + \beta V_{t+1}^j) \right] \\ &= w_t^i + \sum_{j=1}^{n} \left[m_t^{ij} (-\nu \log(m_t^{ij}) - C^{ij} + \beta (V_{t+1}^j - V_{t+1}^i) \right] + \beta V_{t+1}^i \\ &= w_t^i + \sum_{j=1}^{n} \left[m_t^{ij} (\overline{\varepsilon}_t^{ij} - \nu \log(m_t^{ij}) \right] + \beta V_{t+1}^i. \end{aligned}$$

Now, recall from Subsection (B) above that $\log(m^{ij}) = \overline{\varepsilon}_t^{ij}/\nu - \log\left(\sum_{k=1}^{n} \exp(\overline{\varepsilon}^{ik}/\nu)\right)$. This yields:

$$\begin{aligned} V_t^i &= w_t^i + \sum_{j=1}^{n} \left[m_t^{ij} (\nu \log\left(\sum_{k=1}^{n} \exp(\overline{\varepsilon}^{ik}/\nu) \right) \right] + \beta V_{t+1}^i \\ &= w_t^i + \nu \log\left(\sum_{k=1}^{n} \exp(\overline{\varepsilon}^{ik}/\nu) \right) + \beta V_{t+1}^i. \end{aligned}$$

This implies that the option value $\Omega(\overline{\varepsilon}^i)$ can be written as:

$$\Omega(\overline{\varepsilon}^i) = \nu \log\left(\sum_{k=1}^{n} \exp(\overline{\varepsilon}^{ik}/\nu) \right).$$

Alternatively, recalling that $\overline{\varepsilon}^{ii} = 0$, we have:

$$\begin{aligned} \log(m^{ii}) &= 0 - \log\left(\sum_{k=1}^{n} \exp(\overline{\varepsilon}^{ik}/\nu) \right) \\ &= -\log\left(\sum_{k=1}^{n} \exp(\overline{\varepsilon}^{ik}/\nu) \right), \end{aligned}$$

so in equilibrium

$$\Omega(\overline{\varepsilon}^i) = -\nu \log\left(m^{ii} \right).$$

This, then, is (3).

II. Appendix 2: Model with life-cycle features.

A. Basic setup

The economy's workers form a continuum of measure $\overline{L}$. A portion of them, of measure $L_t^{Y,tot}$, are young, and the remainder, of measure $L_t^{O,tot}$, are old. Each period, each young worker will become old with a constant probability λ^Y, and each period, each old worker will drop out of the labor market with probability λ^O, earning a utility of zero from then on. In addition, $\lambda^O L_t^{O,tot}$ new, young workers are added each period.

Each worker at any moment is located in one of the N industries. Denote the number of old, young, and total workers in industry i at the beginning of period t by $L_t^{O,i}$, $L_t^{Y,i}$, and $L_t^i = L_t^{Y,i} + L_t^{O,i}$ respectively. Denote the current allocation vector by $L_t = (L_t^{Y,1}, \ldots, L_t^{Y,n}, L_t^{O,1}, \cdots, L_t^{O,n})$. If a worker, say, $l \in [0, \overline{L}]$, is in industry i at the beginning of t, with age $A \in \{Y, O\}$, she will first learn whether or not she will become old or leave the labor market, effectively immediately; then produce in that industry, collect the market wage $w_t^{A,i}$ for that industry, and then may move to any other industry. In order for the labor market to clear, we must have $w_t^{A,i} = \frac{p^i \partial X^i(L_t^{Y,i}, L_t^{O,i}, s_t)}{\partial L_t^{A,i}}$ at all times, where X^i is the production function for sector i, p^i is the domestic price of sector i's output, and s_t is a state variable following a Markov process.

For the moment assume that all workers have the same educational level. If worker l moves from industry i to industry j, she incurs a cost $C^{A,ij} \geq 0$, which is the same for all workers of age A and all periods, and is publicly known. In addition, if she is in industry i at the end of period t, she collects an idiosyncratic benefit $\varepsilon_{l,t}^i$ from being in that industry. These benefits are independently and identically distributed across individuals, industries, ages, and dates, with density function $f : \Re \longmapsto \Re^+$and cumulative distribution function $F : \Re \longmapsto [0, 1]$. Thus, the full cost for worker l of moving from i to j can be thought of as $\varepsilon_{l,t}^i - \varepsilon_{l,t}^j + C^{A,ij}$. The worker knows the values of the $\varepsilon_{l,t}^i$ for all i before making the period-t moving decision.[1] We adopt the convention that $C^{A,ii} = 0$ for all A, i.

A new worker l entering the labor market at time t can choose the sector in which to locate after learning her realized $\varepsilon_{l,t}$ vector for the period, and pays no entry cost to do so. Once she chooses her sector, say i, she produces there, earns the wage $w_t^{A,i}$, and enjoys her idiosyncratic benefit, $\varepsilon_{l,t}^i$.

All agents have rational expectations and a common constant discount factor $\beta < 1$, and are risk neutral.

B. The key equilibrium condition.

Suppose that we have somehow computed the maximized value to each age-A worker of being in industry i when the labor allocation is L and the state is s. Let $U^{A,i}(L, s, \varepsilon)$ denote this value, which, of course, depends on the worker's realized idiosyncratic shocks. Denote by $V^{A,i}(L, s)$ the expected utility of an A-worker in industry i before learning her realized value of ε and also before learning whether or not she will experience an age transition this period.

Assuming optimizing behavior, i.e., that a worker in industry i will choose to remain at or move to the industry j that offers her the greatest expected benefits, net of moving costs, we can

[1] It is useful to think of the timeline as follows: The worker observes s_t and the vector $\varepsilon_{l,t}$ at the beginning of the period, learns whether or not she will become old or leave the labor market (effective immediately); then, if still in the labor market, produces output and receives the wage, then decides whether or not to move. At the end of the period, if not retired, she enjoys $\varepsilon_{l,t}^j$ in whichever sector j she has landed.

write:[2]

$$U^{A,i}(L_t, s_t, \varepsilon_t) = w_t^{A,i} + \max_j\{\varepsilon_t^j - C^{A,ij} + \beta E_t[V^{A,j}(L_{t+1}, s_{t+1})]\} = w_t^{A,i} + \beta E_t[V^{A,i}(L_{t+1}, s_{t+1})] + \max_j\{\varepsilon_t^{A,j} + \overline{\varepsilon}_t^{A,ij}\} \tag{4}$$

where:

$$\overline{\varepsilon}_t^{A,ij} \equiv \beta E_t[V^{A,j}(L_{t+1}, s_{t+1}) - V^{A,i}(L_{t+1}, s_{t+1})] - C^{A,ij} \tag{5}$$

Note that L_{t+1} is the next-period allocation of labor, derived from L_t and the decision rule, and s_{t+1} is the next-period value of the state, which is a random variable whose distribution is determined by s_t. The expectations in (4) and (5) are taken with respect to s_{t+1} and the possible age transition at time t, conditional on all information available at time t.

Taking the expectation of (4) with respect to the ε vector and the age transition then yields, in the case of a young worker:

$$V^{Y,i}(L_t, s_t) = (1-\lambda^Y)[w_t^{Y,i} + \beta E_t[V^{Y,i}(L_{t+1}, s_{t+1})] + \Omega(\overline{\varepsilon}_t^{Y,i})] + \lambda^Y[w_t^{O,i} + \beta E_t[V^{O,i}(L_{t+1}, s_{t+1})] + \Omega(\overline{\varepsilon}_t^{O,i})], \tag{6}$$

where $\overline{\varepsilon}_t^{A,i} = (\overline{\varepsilon}_t^{A,i1}, ..., \overline{\varepsilon}_t^{A,iN})$ and:

$$\Omega(\overline{\varepsilon}_t^{A,i}) = \sum_{j=1}^{N} \int_{-\infty}^{\infty} (\varepsilon^j + \overline{\varepsilon}_t^{A,ij}) f(\varepsilon^j) \prod_{k \neq j} F(\varepsilon^j + \overline{\varepsilon}_t^{A,ij} - \overline{\varepsilon}_t^{A,ik}) d\varepsilon^j. \tag{7}$$

In the case of an old worker, the parallel equation is:

$$V^{O,i}(L_t, s_t) = (1-\lambda^O)[w_t^{O,i} + \beta E_t[V^{O,i}(L_{t+1}, s_{t+1})] + \Omega(\overline{\varepsilon}_t^{O,i})]. \tag{8}$$

We can write these more compactly by introducing the notation $A = R$ to denote the state of retirement, where $w_t^{R,i} = 0 \forall i, t$, $C^{R,ij} = -\infty \forall i, j, i \neq j$, $V_t^{R,i}(L_t, s_t) = 0 \forall i, t, L_t, s_t$, and, slightly abusing notation, $\overline{\varepsilon}_t^{R,ij} = -\infty \forall i, j, t i \neq j$. Since $\Omega(\overline{\varepsilon}_t^{A,i}) = E_\varepsilon \max_j\{\varepsilon^j + \overline{\varepsilon}_t^{A,ij}\}$ and $\overline{\varepsilon}_t^{A,ii} \equiv 0$, this last condition simply sets $\Omega(\overline{\varepsilon}_t^{R,i}) \equiv 0$. Using this notation, we can write (6) and (8) compactly as:

$$V^{A,i}(L_t, s_t) = E_{A'}[w_t^{A',i} + \beta E_t[V^{A',i}(L_{t+1}, s_{t+1})] + \Omega(\overline{\varepsilon}_t^{A',i})]], \tag{9}$$

where if $A = Y$, A' takes a value of Y with probability $1-\lambda^Y$ and O with probability λ^Y, and if $A = O$, A' takes a value of O with probability $1-\lambda^O$ and R with probability λ^O.

[2]From here on, we drop the worker-specific subscript, l.

Using (9), we can rewrite (5) as:

$$\begin{aligned} C^{A,ij} + \bar{\varepsilon}_t^{A,ij} &= \beta E_t[V^{A,j}(L_{t+1}, s_{t+1}) - V^{A,i}(L_{t+1}, s_{t+1})] \\ &= \beta E_t[w_{t+1}^{A',j} - w_{t+1}^{A',i} + \beta E_{t+1}[V^{A',j}(L_{t+2}, s_{t+2}) - V^{A',i}(L_{t+2}, s_{t+2})] \\ &\quad + \Omega(\bar{\varepsilon}_{t+1}^{A',j}) - \Omega(\bar{\varepsilon}_{t+1}^{A',i})], \text{ or} \end{aligned}$$

$$(10) \qquad C^{A,ij} + \bar{\varepsilon}_t^{A,ij} = \beta E_t[w_{t+1}^{A',j} - w_{t+1}^{A',i} + C^{A',ij} + \bar{\varepsilon}_{t+1}^{A',ij} + \Omega(\bar{\varepsilon}_{t+1}^{A',j}) - \Omega(\bar{\varepsilon}_{t+1}^{A',i})].$$

Here, the left-hand side is evaluated after the date-t age transition has been revealed, so the age A applies through period t to the beginning of period $t+1$. The age A' on the right-hand side is, then, the age for period $t+1$ to the beginning of period $t+2$. Since the value function V is evaluated each period before that period's age transition is revealed, the time-$(t+2)$ value function is conditioned on A'. As before, we adopt the convention that the expectations operator E_t takes expectations over A' as well as the other variables.

C. The estimating equation.

Let $m_t^{A,ij}$ be the fraction of the age-A labor force in industry i at time t that chooses to move to industry j, i.e., the gross flow from i to j. If we assume, as in the main model, that the idiosyncratic shocks follow an extreme-value distribution, then following the algebra of Appendix 1, amending slightly to control for age, we obtain:

$$(11) \qquad \bar{\varepsilon}_t^{A,ij} \equiv \beta E_t[V_{t+1}^{A,j} - V_{t+1}^{A,i}] - C^{A,ij} = \nu[\ln m_t^{A,ij} - \ln m_t^{A,ii}]$$

and:

$$(12) \qquad \Omega(\bar{\varepsilon}_t^{A,i}) = -\nu \ln m_t^{A,ii}.$$

Substituting from (11) and (12) into (10) and rearranging, we get the following conditional moment condition:

$$(13) \qquad E_t\left[\frac{\beta}{\nu}(w_{t+1}^{A',j} - w_{t+1}^{A',i}) + \beta(\ln m_{t+1}^{A',ij} - \ln m_{t+1}^{A',jj}) + \frac{(\beta C^{A',ij} - C^{A,ij})}{\nu} - (\ln m_t^{A,ij} - \ln m_t^{A,ii})\right] = 0.$$

In the case of a young worker this amounts to:

$$\begin{aligned} &E_t[\frac{\beta}{\nu}[((1-\lambda^Y)w_{t+1}^{Y,j} + \lambda^Y w_{t+1}^{O,j}) - ((1-\lambda^Y)w_{t+1}^{Y,i} + \lambda^Y w_{t+1}^{O,i})] \\ &+\beta[(1-\lambda^Y)\ln m_{t+1}^{Y,ij} + \lambda^Y \ln m_{t+1}^{O,ij} - ((1-\lambda^Y)\ln m_{t+1}^{Y,jj} + \lambda^Y \ln m_{t+1}^{O,jj})] \\ &+\frac{(\beta(1-\lambda^Y) - 1)C^{Y,ij} + \beta\lambda^Y C^{O,ij}}{\nu} \\ (14) \qquad &-(\ln m_t^{Y,ij} - \ln m_t^{Y,ii})] = 0. \end{aligned}$$

In the case of an old worker this amounts to:

$$E_t[\frac{\beta}{\nu}(1-\lambda^O)[(w^{O,j}_{t+1})-(w^{O,i}_{t+1})]+\beta(1-\lambda^O)[(\ln m^{O,ij}_{t+1}-\ln m^{O,jj}_{t+1})] + \frac{(\beta(1-\lambda^O)-1)C^{O,ij}}{\nu} - (\ln m^{O,ij}_t - \ln m^{O,ii}_t)] = 0. \tag{15}$$

Conditions (14) and (15) can then be used together to estimate the moving cost parameters. Once we have decided on a cutoff age to separate "young" from "old," we set λ^Y and λ^O so that the average length of each state is equal to the actual duration of the state. In practice, we define workers aged 25 to 44 as young, and workers 45 to 65 as old, so we set $\lambda^Y = \lambda^0 = 0.05$, to make the duration of each state 20 years.[3] Thus, the only parameters to estimate are $C^{Y,ij}$, $C^{O,ij}$, ν, and, in principle, β.

It is now trivial to add different human capital types. Assume that each worker at the beginning of her productive life is either college educated or not college educated; that this is the only human-capital distinction that matters; and that workers never switch between those two categories. Then (14) and (15) apply conditional on educational status, and we have four common cost parameters, $C^{A,E,ij}$, to estimate, one for each age-education state, where E stands for education level. In principle, we could estimate the ν parameter separately for each category as well, but degrees of freedom issues have discouraged us from doing so.

Note that for each educational class, (14) and (15) are a system of two equations with common parameters, and taking both classes we have four equations with a common parameter of ν. We therefore use the GMM method adapted for systems of equations with unknown heteroskedasticity, as in Greene (2000, pp. 696-98).

D. Simulation.

We need to specify a production function for each sector, which must have all four types of labor as well as capital as arguments. To reduce the dimensionality of the problem, we assume that there is a CES aggregator for labor across ages:

$$\widetilde{L}^{E,i} \equiv (\alpha^E(L^{Y,E,i})^{\rho^E} + (1-\alpha^E)(L^{O,E,i})^{\rho^E})^{\frac{1}{\rho^E}}, \tag{16}$$

where $\widetilde{L}^{E,i}$ is the effective amount of labor of educational level E in sector i, and α^E and ρ^E are positive parameters.

$$y^i_t = \psi^i\left(\alpha^i(\widetilde{L}^{N,i}_t)^{\rho^i} + (1-\alpha^i)(\widetilde{\alpha}^i(\widetilde{L}^{C,i})^{\widetilde{\rho}^i} + (1-\widetilde{\alpha}^i)(K^i)^{\widetilde{\rho}^i})^{\frac{\rho^i}{\widetilde{\rho}^i}}\right)^{\frac{1}{\rho^i}}, \tag{17}$$

where y^i_t is the output for sector i in period t, K^i is sector-i's capital stock, and $\alpha^i \in [0,1]$, $\widetilde{\alpha}^i \in [0,1]$, $\rho^i < 1$, $\widetilde{\rho}^i < 1$ and $\psi^i > 0$ are parameters.

[3]Strictly speaking, this creates a problem because it implies equal numbers of young and old in the steady state, but empircally with this definition of young and old there are considerably more young workers in the economy. This could be remedied, in principle, by raising the threshold above 45, which would increase the length of time spent while young and thus lowering λ^Y and at the same time lowering the length of time spent old and thus raising λ^O. This would, therefore, imply a lower steady state fraction of the population classified as old, and with the appropriate choice of threshold, the proportions in the data could be matched. However, with our data, older workers are scarce and this would make it difficult to estimate the parameters for older workers.

As in the homogeneous-labor case, we choose parameters to provide a plausible illustrative example to minimize a loss function. For the period of our data, we have $L_t^{E,A,i}$ for all E, A, i and t and so for any choice of parameter values can generate the wages, share of labor in unit cost for each sector, and share of each sector's output in GDP. For each year, we get the sum of the squared deviation of these values from the actual values in the data (labor shares and GDP shares from the BEA), and choose the parameters to minimize the sum of those squared deviations over all years.

TABLE 1: PARAMETERS FOR SIMULATION OF HETEROGENOUS-WORKER MODEL.

Economy-wide parameters.								
ρ^N	0.968							
ρ^C	0.99							
α^N	0.451							
α^C	0.479							
Sector-specific parameters.								
	α^i	$\widetilde{\alpha}^i$	ρ^i	$\widetilde{\rho}^i$	ψ^i	Consumer share.	Domestic price.	World price.
Agric/Min	0.3102	0.1038	0.2596	0.1070	0.5912	0.07	1	1
Const	0.5265	0.4277	0.4356	0.4029	1.7089	0.3	1	1*
Manuf	0.1917	0.0841	0.1066	0.01	2.2973	0.3	1	0.7
Trans/Util	0.2205	0.3874	0.01	0.4348	1.7206	0.08	1	1*
Trade	0.4473	0.3532	0.4281	0.2849	1.9198	0	1	1*
Service	0.3547	0.4265	0.99	0.5937	4.6313	0.25	1	1

(Note:* Under the second simulation specification, the sectors marked with an asterisk are non-traded, so they have no world price.)

We use the same algorithm for solving the perfect-foresight equilibrium as in the basic model, laid out in Artuç, Chaudhuri and McLaren (2008). One note that we should make concerns the treatment of new workers. New workers are 29% college-educated and 71% non-college educated, and are allowed to choose their sector of first employment to solve:

$$max_i[w^{E,Y,i} + \varepsilon_t^{l,i} + \beta E_t[V^{E,Y,i}(L_{t+1}, s_{t+1})]], \tag{18}$$

where $\varepsilon_t^{l,i}$ is new worker l's realized idiosyncratic shock. The new worker pays no moving cost because she is not changing sectors, simply choosing her first sector. This implies an allocation of new workers as follows:

$$m_t^{E,Y.0i} = \frac{exp(\frac{\beta}{\nu}V^{E,Y,i}(L_{t+1}, s_{t+1}))}{\sum_j exp(\frac{\beta}{\nu}V^{E,Y,j}(L_{t+1}, s_{t+1}))}, \tag{19}$$

where $m_t^{E,Y.0i}$ denotes the fraction of new entrant of educational type E who choose sector i as first sector of employment.

Appendix 3: Unobserved Worker Heterogeneity.

In this version we have two types of workers, indexed by $A = \{1, 2\}$, who differ only in their (common) moving cost C^A. Any type-1 worker can become a type-2 worker at any time and vice-versa, and the probability of switching from type A to the other type is λ_A. We write the Euler equations for both types, and then from them derive a condition that must hold in the limit as $C^2 \to \infty$.

Define $X_t^{A,ij} = \nu\left(\ln m_t^{A,ij} - \ln m_t^{A,jj}\right) + C^A$, $Y_t^{A,ij} = \nu\left(\ln m_t^{A,ij} - \ln m_t^{A,ii}\right) + C^A$, and $\Delta w_t^{ij} = w_t^j - w_t^i$ The Euler equations for the two types are then:

$$
\begin{aligned}
Y_t^1 &= E_t[(1-\lambda_1)\left[\beta\Delta w_{t+1} + \beta X_{t+1}^1\right] + \lambda_1\left[\beta\Delta w_{t+1} + \beta X_{t+1}^2\right]],\\
Y_t^2 &= E_t\,(1-\lambda_2)\left[\beta\Delta w_{t+1} + \beta X_{t+1}^2\right] + \lambda_2\left[\beta\Delta w_{t+1} + \beta X_{t+1}^1\right]].
\end{aligned}
$$

We can evaluate the Euler equation of type 1 workers at time $t-1$ and t, then re-arrange them such that:

$$
\begin{aligned}
E_t X_{t+1}^2 &= \frac{1}{\beta\lambda_1} E_t\left\{Y_t^1 - \beta\Delta w_{t+1} - (1-\lambda_1)\beta X_{t+1}^1\right\},\\
E_{t-1} X_t^2 &= \frac{1}{\beta\lambda_1} E_{t-1}\left\{Y_{t-1}^1 - \beta\Delta w_t - (1-\lambda_1)\beta X_t^1\right\}.
\end{aligned}
$$

Type 2 workers' Euler condition can be re-arranged in the following way:

$$
E_t Y_t^2 - (1-\lambda_2)\beta E_t X_{t+1}^2 = E_t\left\{\beta\Delta w_{t+1} + \lambda_2\beta E_t X_{t+1}^1\right\}.
$$

Note that for suffiently large C^2, $X_t^2 = Y_t^2$ since $\ln m_t^{2,ii}, \ln m_t^{2,jj} \to 0$ as $C^2 \to \infty$. Finally, we can plug X_{t+1}^2 and X_t^2 from type 1 workers' Euler equations into the equation above. This gives an equation in variables dated at time $t-1$, t, and $t+1$. Shifting the time index forward one period for convenience gives:

$$
\begin{aligned}
\ln m_t^{1,ij} - \ln m_t^{1,ii} &= (1-\lambda_1)\beta E_t[\ln m_{t+1}^{1,ij} - \ln m_{t+1}^{1,jj}] + (1-\lambda_2)\beta E_t[\ln m_{t+1}^{1,ij} - \ln m_{t+1}^{1,ii}]\\
&\quad + (\lambda_1+\lambda_2-1)\beta^2 E_t[\ln m_{t+2}^{1,ij} - \ln m_{t+2}^{1,jj}]\\
&\quad + \frac{\beta}{\nu} E_t\left[w_{t+1}^j - w_{t+1}^i\right] + (\lambda_1+\lambda_2-1)\frac{\beta^2}{\nu} E_t\left[w_{t+2}^j - w_{t+2}^i\right]\\
&\quad + \frac{C^1}{\nu}\left\{(\lambda_1+\lambda_2-1)\beta^2 + (2-\lambda_1-\lambda_2)\beta - 1\right\}.
\end{aligned}
$$

This is, then, an estimating equation, which can be estimated by GMM in a manner completely analogous to estimation of the main model. Since we let C^2 go to infinity, we can drop the superscript and denote the common moving cost for type 1 as C.

Finally, we comment on how we choose α, λ_1 and λ_2 for our example. It is possible to calibrate these parameters from panel data where it is possible to see the history of each worker. Let x be the rate of gross flow in the economy of type 1 (mobile) workers. The observed gross flows will be $(1-\alpha)\,x$. If a worker has changed her sector (which means that she is type 1) at time $t-1$, her probability of moving at t is $(1-\lambda_1)\,x$. If a worker has changed her sector at time $t-2$, her probability of moving at time t is $(1-\lambda_1)^2\,x+\lambda_1\lambda_2 x$. Finally to have a constant number of immobile workers over time we must have $\alpha\lambda_2=(1-\alpha)\,\lambda_1$. This is a system of four equations with four unknowns. We do this exercise with the NLSY just to show that it is feasible and find that $\alpha=0.75$, $\lambda_1=0.44$ and $\lambda_2=0.15$.